

CLAIMS

What is claimed is:

- 1 1. A method for performing database recovery after a crash of an instance of a
2 database, wherein multiple transactions were active when the instance
3 crashed, the method comprising the steps of:
4 identifying a plurality of dead transactions;
5 determining statistical data about said plurality of dead transactions;
6 determining that a particular number of recovery servers should be used to recover
7 said plurality of dead transactions based on the statistical data; and
8 recovering said plurality of dead transactions using said particular number of
9 recovery servers.
- 1 2. The method of Claim 1, wherein the step of recovering said plurality of dead
2 transactions is performed by executing the particular number of recovery
3 servers in parallel.
- 1 3. The method of Claim 1, wherein:
2 the step of identifying the plurality of dead transactions includes the step of
3 maintaining a working list, wherein the working list identifies a list of
4 dead transactions for which recovery will be attempted; and
5 the step of determining statistical data includes the step of determining statistical
6 data based on the list of dead transactions.
- 1 4. The method of Claim 3, wherein the step of maintaining a working list
2 comprises the steps of:

3 locating a rollback segment, wherein the rollback segment contains a transaction
4 table that contains entries associated with dead transactions;
5 scanning the transaction table to identify the dead transactions; and
6 storing the identity of the dead transactions in the working list.

1 5. The method of Claim 4, wherein:

2 the method further comprises the step of maintaining a block count, wherein the
3 block count identifies the number of undo blocks that are associated with a
4 particular transaction; and
5 the step of determining statistical data includes the step of determining a total
6 number of undo blocks that need to be recovered, wherein the total
7 number of undo blocks is based on the block count associated with the
8 dead transactions identified in the working list.

1 6. The method of Claim 4, wherein the step of determining statistical data
2 includes the step of determining statistical data based on the number of dead
3 transactions that are identified in the working list.

1 7. The method of Claim 1, wherein the step of determining that the particular
2 number of recovery servers should be used includes the step of determining
3 that the particular number of recovery servers should be used based on a
4 max_parallelism threshold value, wherein the max_parallelism threshold
5 value provides an upper limit for the number of recovery servers to be used.

1 8. The method of Claim 7, further comprises the step of determining the
2 max_parallelism threshold value based on a user input value.

- 1 9. The method of Claim 1, further comprising the steps of:
- 2 identifying a rollback segment that was previously owned by the crashed instance
- 3 at the time of its crash; and
- 4 the crashed instance reacquiring ownership of the rollback segment after the
- 5 crashed instance is restarted.
- 1 10. The method of Claim 9, wherein the step of reacquiring ownership of the
- 2 rollback segment includes the steps of:
- 3 identifying an instance that currently owns the rollback segment that was
- 4 previously owned by the crashed instance at the time of its crash;
- 5 requesting the instance to release ownership of the rollback segment; and
- 6 the instance releasing ownership of the rollback segment in response to the
- 7 request.
- 1 11. The method of Claim 1, further comprises the steps of:
- 2 identifying a rollback segment that is unowned, wherein the unowned rollback
- 3 segment is not currently associated with any instance of the database; and
- 4 associating the unowned rollback segment with the crashed instance, wherein
- 5 associating the unowned rollback segment with the crashed node causes
- 6 the rollback segment to be owned by the crashed instance.
- 1 12. The method of Claim 1, wherein the step of recovering the plurality of dead
- 2 transactions comprises the steps of:
- 3 maintaining a working list, wherein the working list identifies a list of dead
- 4 transactions for which recovery will be attempted;
- 5 selecting a dead transaction from the working list;

6 acquiring a rollback segment lock on a rollback segment, wherein the rollback
7 segment is associated with a transaction table that contains an entry that
8 corresponds to the dead transaction;
9 acquiring a transaction lock on a chain of undo, wherein the chain of undo contain
10 change information associated with the dead transaction;
11 determining whether the dead transaction still needs to be recovered; and
12 if the dead transaction still needs to be recovered, assigning the dead transaction
13 to a recovery server.

1 13. The method of Claim 12, wherein the step of acquiring the transaction lock
2 includes the step of a coordinator acquiring the transaction lock.

1 14. The method of Claim 13, wherein the method further comprises the steps of:

2 upon completing the recovery of the dead transaction, the recovery server
3 signaling the coordinator to indicate it has completed the recovery of the
4 dead transaction; and
5 upon receiving the signal from the recovery server, the coordinator releasing its
6 lock on the transaction.

1 15. The method of Claim 1, wherein the step of recovering the plurality of dead
2 transactions using the particular number of recovery servers includes the
3 steps of:

4 assigning two or more dead transactions to a recovery server;
5 associating a time slice value with the recovery server, wherein the time slice
6 value is used by the recovery server to promote fairness during recovery of
7 the two or more dead transactions; and
8 recovering the two or more dead transactions using the time slice value.

- 1 16. A computer-readable medium carrying one or more sequences of one or
2 more instructions for performing database recovery after a crash of an
3 instance of a database, wherein multiple transactions were active when the
4 instance crashed, wherein the execution of the one or more sequences of one
5 or more instructions by one or more processors causes the one or more
6 processors to perform the steps of:
- 7 identifying a plurality of dead transactions;
8 determining statistical data about said plurality of dead transactions;
9 determining that a particular number of recovery servers should be used to recover
10 said plurality of dead transactions based on the statistical data; and
11 recovering said plurality of dead transactions using said particular number of
12 recovery servers.
- 1 17. The computer-readable medium of Claim 16, wherein the step of recovering
2 said plurality of dead transactions is performed by executing the particular
3 number of recovery servers in parallel.
- 1 18. The computer-readable medium of Claim 16, wherein:
- 2 the step of identifying the plurality of dead transactions includes the step of
3 maintaining a working list, wherein the working list identifies a list of
4 dead transactions for which recovery will be attempted; and
5 the step of determining statistical data includes the step of determining statistical
6 data based on the list of dead transactions.
- 1 19. The computer-readable medium of Claim 18, wherein the step of
2 maintaining a working list comprises the steps of:

3 locating a rollback segment, wherein the rollback segment contains a transaction
4 table that contains entries associated with dead transactions;
5 scanning the transaction table to identify the dead transactions; and
6 storing the identity of the dead transactions in the working list.

1 20. The computer-readable medium of Claim 19, wherein:

2 the computer-readable medium further comprises instructions for performing the
3 step of maintaining a block count, wherein the block count identifies the
4 number of undo blocks that are associated with a particular transaction;
5 and
6 the step of determining statistical data includes the step of determining a total
7 number of undo blocks that need to be recovered, wherein the total
8 number of undo blocks is based on the block count associated with the
9 dead transactions identified in the working list.

1 21. The computer-readable medium of Claim 19, wherein the step of
2 determining statistical data includes the step of determining statistical data
3 based on the number of dead transactions that are identified in the working
4 list.

1 22. The computer-readable medium of Claim 16, wherein the step of
2 determining that the particular number of recovery servers should be used
3 includes the step of determining that the particular number of recovery
4 servers should be used based on a max_parallelism threshold value, wherein
5 the max_parallelism threshold value provides an upper limit for the number
6 of recovery servers to be used.

- 1 23. The computer-readable medium of Claim 22, further comprises instructions
2 for performing the step of determining the max_parallelism threshold value
3 based on a user input value.
- 1 24. The computer-readable medium of Claim 16, further comprising instructions
2 for performing the steps of:
3 identifying a rollback segment that was previously owned by the crashed instance
4 at the time of its crash; and
5 the crashed instance reacquiring ownership of the rollback segment after the
6 crashed instance is restarted.
- 1 25. The computer-readable medium of Claim 24, wherein the step of
2 reacquiring ownership of the rollback segment includes the steps of:
3 identifying an instance that currently owns the rollback segment that was
4 previously owned by the crashed instance at the time of its crash;
5 requesting the instance to release ownership of the rollback segment; and
6 the instance releasing ownership of the rollback segment in response to the
7 request.
- 1 26. The computer-readable medium of Claim 16, further comprises instructions
2 for performing the steps of:
3 identifying a rollback segment that is unowned, wherein the unowned rollback
4 segment is not currently associated with any instance of the database; and
5 associating the unowned rollback segment with the crashed instance, wherein
6 associating the unowned rollback segment with the crashed node causes
7 the rollback segment to be owned by the crashed instance.

1 27. The computer-readable medium of Claim 16, wherein the step of recovering
2 the plurality of dead transactions comprises the steps of:

3 maintaining a working list, wherein the working list identifies a list of dead
4 transactions for which recovery will be attempted;
5 selecting a dead transaction from the working list;
6 acquiring a rollback segment lock on a rollback segment, wherein the rollback
7 segment is associated with a transaction table that contains an entry that
8 corresponds to the dead transaction;
9 acquiring a transaction lock on a chain of undo, wherein the chain of undo contain
10 change information associated with the dead transaction;
11 determining whether the dead transaction still needs to be recovered; and
12 if the dead transaction still needs to be recovered, assigning the dead transaction
13 to a recovery server.

1 28. The computer-readable medium of Claim 27, wherein the step of acquiring
2 the transaction lock includes the step of a coordinator acquiring the
3 transaction lock.

1 29. The computer-readable medium of Claim 28, wherein the computer-
2 readable medium further comprises instructions for performing the steps of:

3 upon completing the recovery of the dead transaction, the recovery server
4 signaling the coordinator to indicate it has completed the recovery of the
5 dead transaction; and
6 upon receiving the signal from the recovery server, the coordinator releasing its
7 lock on the transaction.

- 1 30. The computer-readable medium of Claim 16, wherein the step of recovering
2 the plurality of dead transactions using the particular number of recovery
3 servers includes the steps of:
4 assigning two or more dead transactions to a recovery server;
5 associating a time slice value with the recovery server, wherein the time slice
6 value is used by the recovery server to promote fairness during recovery of
7 the two or more dead transactions; and
8 recovering the two or more dead transactions using the time slice value.
- 1 31. A system for performing database recovery after a crash of an instance of a
2 database, wherein multiple transactions were active when the instance
3 crashed, the system comprising:
4 a memory;
5 one or more processors coupled to the memory; and
6 a set of computer instructions contained in the memory, the set of computer
7 instructions including computer instructions which when executed by one
8 or more processors, cause the one or more processors to perform the steps
9 of:
10 identifying a plurality of dead transactions;
11 determining statistical data about said plurality of dead transactions;
12 determining that a particular number of recovery servers should be used to
13 recover said plurality of dead transactions based on the statistical
14 data; and
15 recovering said plurality of dead transactions using said particular number
16 of recovery servers.

1 32. The system of Claim 31, wherein the step of recovering said plurality of
2 dead transactions is performed by executing the particular number of
3 recovery servers in parallel.

1 33. The system of Claim 31, wherein:
2 the step of identifying the plurality of dead transactions includes the step of
3 maintaining a working list, wherein the working list identifies a list of
4 dead transactions for which recovery will be attempted; and
5 the step of determining statistical data includes the step of determining statistical
6 data based on the list of dead transactions.

1 34. The system of Claim 33, wherein the step of maintaining a working list
2 comprises the steps of:
3 locating a rollback segment, wherein the rollback segment contains a transaction
4 table that contains entries associated with dead transactions;
5 scanning the transaction table to identify the dead transactions; and
6 storing the identity of the dead transactions in the working list.

1 35. The system of Claim 34, wherein:
2 the system further comprises the step of maintaining a block count, wherein the
3 block count identifies the number of undo blocks that are associated with a
4 particular transaction; and
5 the step of determining statistical data includes the step of determining a total
6 number of undo blocks that need to be recovered, wherein the total
7 number of undo blocks is based on the block count associated with the
8 dead transactions identified in the working list.

- 1 36. The system of Claim 31, wherein the step of determining that the particular
2 number of recovery servers should be used includes the step of determining
3 that the particular number of recovery servers should be used based on a
4 max_parallelism threshold value, wherein the max_parallelism threshold
5 value provides an upper limit for the number of recovery servers to be used.
- 1 37. The system of Claim 31, further comprising the steps of:
2 identifying a rollback segment that was previously owned by the crashed instance
3 at the time of its crash; and
4 the crashed instance reacquiring ownership of the rollback segment after the
5 crashed instance is restarted.
- 1 38. The system of Claim 31, further comprises the steps of:
2 identifying a rollback segment that is unowned, wherein the unowned rollback
3 segment is not currently associated with any instance of the database; and
4 associating the unowned rollback segment with the crashed instance, wherein
5 associating the unowned rollback segment with the crashed node causes
6 the rollback segment to be owned by the crashed instance.
- 1 39. The system of Claim 31, wherein the step of recovering the plurality of dead
2 transactions comprises the steps of:
3 maintaining a working list, wherein the working list identifies a list of dead
4 transactions for which recovery will be attempted;
5 selecting a dead transaction from the working list;

6 acquiring a rollback segment lock on a rollback segment, wherein the rollback
7 segment is associated with a transaction table that contains an entry that
8 corresponds to the dead transaction;
9 acquiring a transaction lock on a chain of undo, wherein the chain of undo contain
10 change information associated with the dead transaction;
11 determining whether the dead transaction still needs to be recovered; and
12 if the dead transaction still needs to be recovered, assigning the dead transaction
13 to a recovery server.

1 40. The system of Claim 31, wherein the step of recovering the plurality of dead
2 transactions using the particular number of recovery servers includes the
3 steps of:
4 assigning two or more dead transactions to a recovery server;
5 associating a time slice value with the recovery server, wherein the time slice
6 value is used by the recovery server to promote fairness during recovery of
7 the two or more dead transactions; and
8 recovering the two or more dead transactions using the time slice value.